GESTION QUANTITATIVE

SOMMAIRE

INTRODUCTION……2

I. Analyses descriptives des données : première étude exploratoire……3

II. Présentation la théorie de la tarification……7

III. Modélisation de la fréquence des sinistres. ……7

IV. Modélisation de du cout des sinistres. ……10

V. Détermination de la prime pure. ……12

VI. Analyse des résultats et conclusion. ……13

VII. Annexe……13

Introduction

Pour évaluer le montant des cotisations qu'un souscripteur devra verser, les compagnies d'assurance auto se basent sur de nombreux critères, liés à la fois au véhicule et à son conducteur. Cette évaluation repose sur la charge que l'assureur devra couvrir pour indemniser son assuré en cas de sinistre, de manière à ce que la prime reflète adéquatement le risque à prendre. On constate que l'estimation de ces montants dépend de plusieurs critères, que l'on appelle critères de tarification. Dans le cadre de notre étude sur la souscription, on retrouve notamment les critères suivants :

- **Critères liés à l'assuré** : par exemple, l'âge pour un particulier.
- **Critères liés au bien assuré (véhicule)** : marque, puissance, âge du véhicule, type de carburant (diesel/essence), etc.
- **Critères géographiques** : zones et régions de circulation habituelle, densité de circulation, etc.

En effet, l'élaboration d'un tarif en assurance IARD (auto, MRH, construction, etc.) s'appuie traditionnellement sur l'analyse de la prime pure dans le cadre d'un modèle *fréquence × coût*, où l'effet des variables explicatives sur le niveau de risque est modélisé via des modèles de régression de type GLM (Modèles Linéaires Généralisés). Ces dernières années, l'amélioration des performances informatiques a conduit à un intérêt croissant pour des approches alternatives, non paramétriques ou semi-paramétriques, qui permettent de contourner certaines limitations des modèles de régression classiques.

Problématique

La probabilité de survenance d'un sinistre varie selon les critères, et le coût du sinistre influe directement sur la prime pure. Dans ce contexte, comment déterminer une tarification adaptée à partir de la base de données d'une compagnie d'assurance dans un pays donné ? La question centrale est : quels critères doivent être retenus dans l'équation de la prime pure pour permettre une estimation précise et justifiée du montant que l'assuré devra payer ?

Présentation de l'étude

L'objectif de notre étude est de construire un modèle permettant d'évaluer au plus près la prime pure, en considérant à la fois la fréquence des sinistres et leur coût. Nous disposons de deux bases de données :

1. **Base fréquence (CASdatasets)** : contient des variables explicatives fournissant des informations sur le conducteur et le véhicule.
2. **Base coût (freMTPLsev)** : contient les informations relatives au coût des sinistres.

Ces deux bases peuvent être reliées via une clé primaire ("PolicyID"). Pour constituer une base combinée contenant le nombre de sinistres et les coûts associés, la fonction *merge* est particulièrement utile dans la phase de modélisation du coût.

Le but est d'obtenir, à partir de ces données, une équation du coût et de la fréquence en fonction des variables de tarification jugées significatives pour le modèle. Pour ce faire, nous utiliserons les Modèles Linéaires Généralisés (GLM) et procéderons préalablement à une segmentation des données.

Notre étude se déroulera selon le plan suivant :

1. Analyse descriptive des données : première exploration.
2. Présentation de la théorie de la tarification.
3. Modélisation de la fréquence des sinistres.
4. Modélisation du coût des sinistres.
5. Détermination de la prime pure.
6. Analyse des résultats et conclusion.

I. Analyse descriptive des données

Dans le cadre de notre étude, nous travaillerons sur deux bases de données : la base de fréquence des sinistres et la base du coût moyen des sinistres, en utilisant le logiciel R.

Traitement des données : À l'aide des fonctions *dim()*, *names()* et *str()*, on observe que la base fréquence contient 10 variables et 412 944 observations, tandis que la base coût contient 2 variables et 16 181 observations.

```
> str(data_frq)
'data.frame':   413169 obs. of  10 variables:
 $ PolicyID : Factor w/ 413169 levels "1","2","3","4",..: 1 2 3 4 5 6 7 8 9 $
 $ ClaimNb  : int  0 0 0 0 0 0 0 0 0 0 ...
 $ Exposure : num  0.09 0.84 0.52 0.45 0.15 0.75 0.81 0.05 0.76 0.34 ...
 $ Power    : Factor w/ 12 levels "d","e","f","g",..: 4 4 3 3 4 4 1 1 1 6 ...
 $ CarAge   : int  0 0 2 2 0 0 1 0 9 0 ...
 $ DriverAge: int  46 46 38 38 41 41 27 27 23 44 ...
 $ Brand    : Factor w/ 7 levels "Fiat","Japanese (except Nissan) or Korean"$
 $ Gas      : Factor w/ 2 levels "Diesel","Regular": 1 1 2 2 1 1 2 2 2 2 ...
 $ Region   : Factor w/ 10 levels "Aquitaine","Basse-Normandie",..: 1 1 8 8 $
 $ Density  : int  76 76 3003 3003 60 60 695 695 7887 27000 ...
```

```
> names(data_frq)
 [1] "PolicyID"  "ClaimNb"   "Exposure"  "Power"     "CarAge"    "DriverAge"
 [7] "Brand"     "Gas"       "Region"    "Density"
```

```
> str(data_cout)
'data.frame':   16181 obs. of  2 variables:
 $ PolicyID   : int  63987 310037 314463 318713 309380 309380 318738 305914 313$
 $ ClaimAmount: int  1172 1905 1150 1220 55077 7593 1176 1202 1203 1232 ...
>
```

On observe que les variables dans la base de la fréquence, il existe :

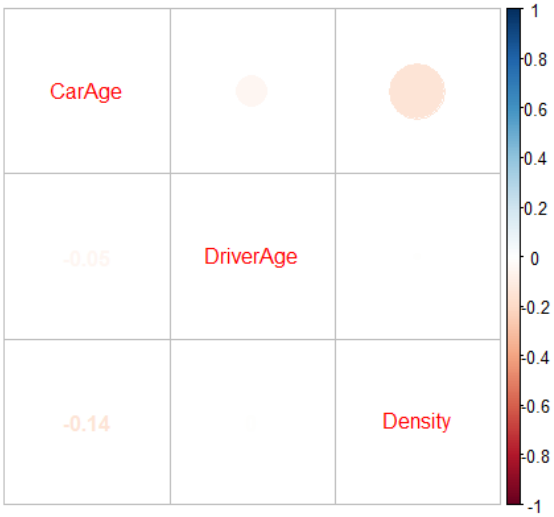
- ⇨ Variable quantitative : DriveAge, CarAge, Density.
- ⇨ Variable qualitative : Power, Brand, Gas, Region.
- ⇨ Clé primaire : PolicyID.
- ⇨ Variable expliquée et l'existence du sinistre : ClaimNb et Exposure.

Et dans la base du coût, on a des variables :

- ⇨ La clé primaire : PolicyID.
- ⇨ Le coût du sinistre : ClaimAmount.

➢ Corrélation

```
> mcor <- cor(data_frq[,var_quanti])
> mcor
               CarAge     DriverAge       Density
CarAge      1.00000000 -0.046413527 -0.142318327
DriverAge -0.04641353  1.000000000 -0.001692481
Density    -0.14231833 -0.001692481  1.000000000
```



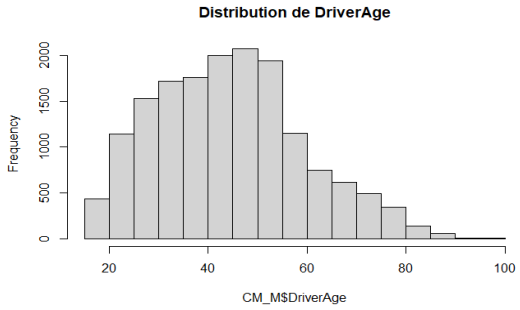
Résultats de la corrélation

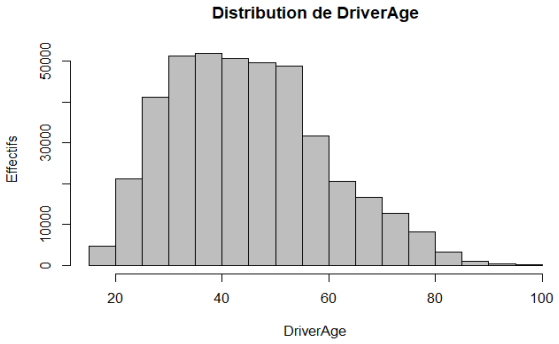
L'analyse de corrélation montre une relation négative entre l'âge du véhicule et la densité de population ainsi que le nombre de véhicules. Autrement dit, plus la densité de population est élevée, plus les véhicules ont tendance à être récents.

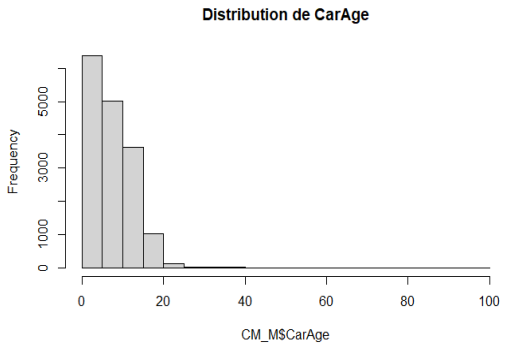
Histogrammes

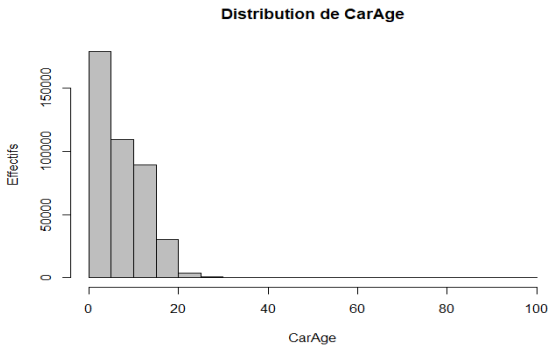
Nous présentons des histogrammes pour plusieurs variables : âge du conducteur, âge du véhicule, densité de population et coût des sinistres.

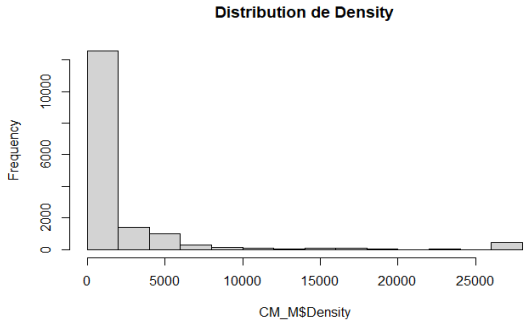
- Les graphiques situés à gauche montrent les **distributions en fonction de la fréquence des sinistres**.
- Les graphiques situés à droite représentent les **distributions en fonction de la probabilité d'occurrence d'un sinistre**.

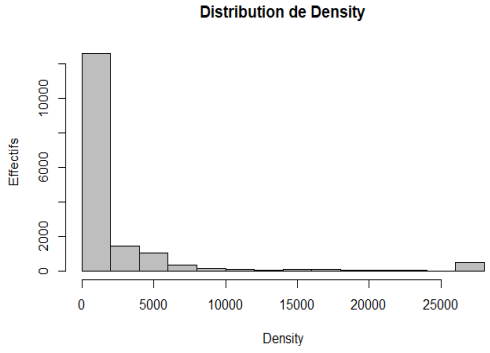


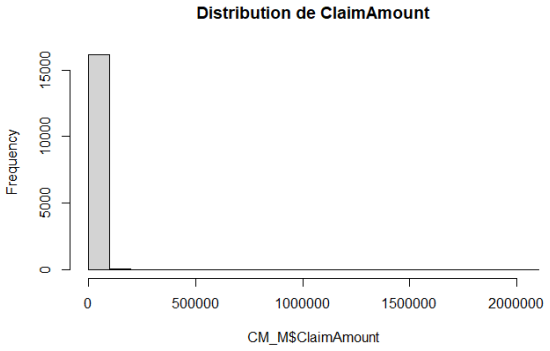


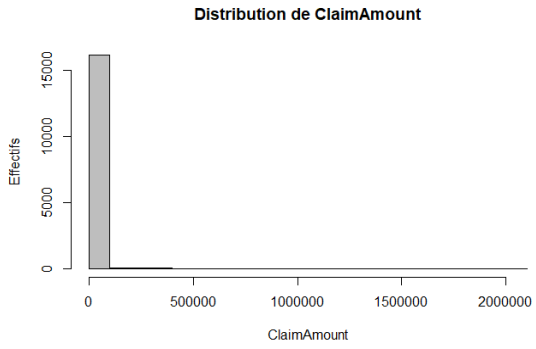


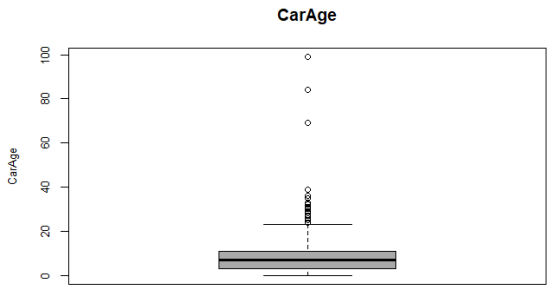


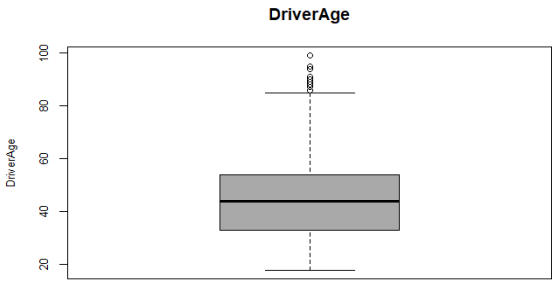


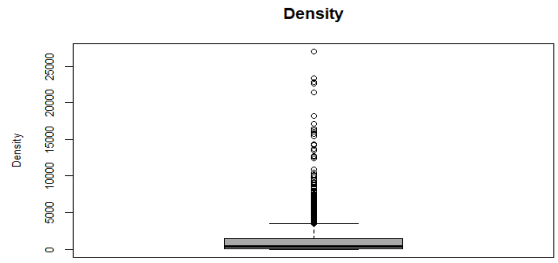


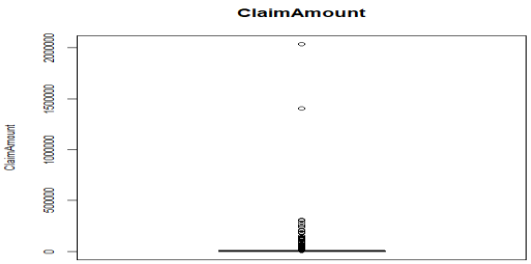












- Calcul de la fréquence

```
> calcul_freq=sum(data_frq$ClaimNb)/sum(data_frq$Exposure)
> print(calcul_freq)
[1] 0.06979859
```

La fréquence est égale à 0,06979859.

Calcul du cout

```
> calcul_cout=sum(data_cout$ClaimAmount)/sum(data_frq$ClaimNb)
> print(calcul_cout)
[1] 2129.972
```

Le cout est égal à 2129,972.

Ensuite on peut calculer la Prime pure, on a :

$$Prime\ pure = Fréquence \times Coût\ moyen$$

$$Prime\ pure = 0{,}06979859 \times 2129{,}972 = 148{,}6690423$$

Cette prime pure permet d'expliquer la structure de la prime pure observée dans notre base de données générale. Toutefois, ce calcul n'est pas nécessairement adapté à l'ensemble des assureurs, car les informations disponibles varient selon les profils d'assurés.
La question qui se pose alors est la suivante : **comment établir une tarification spécifique et équitable pour chaque assuré** en tenant compte de ses caractéristiques propres ?
Nous aborderons cette problématique dans la section suivante.

II. Présentation la théorie de la tarification.

Dans une entreprise d'assurance, la fréquence annualisée est le nombre de sinistres divisé par l'exposition (correspondant au nombre d'années d'une entreprise d'assurance dans un paye). La plupart des contrats étant annuels, on ramènera le nombre de sinistres à une exposition annuelle lors du calcul de la prime, on notera N (ClaimNb/ Exposition) le nombre du sinistre. Durant la période d'exposition, on notera A_i les coûts du sinistres (ClaimAmount), c'est les indemnités versées par l'assureur à l'assuré. La variable « ClaimNb », numéro « 1 » ça signifie qu'il existe l'exposition du sinistre, et le numéro « 0 » ça signifie qu'il n'existe pas du sinistre. C'est la charge totale d'une entreprise d'assurance.

$$S= A_1 + \dots + A_N = \sum_{i=1}^{N} A_i.$$

La prime pure égale : $E(S) = E(N) * E(Y_i)$

E(N) : Nombre du sinistre total ($\sum_{i=1}^{N}$(ClaimNb/ Exposition))

$E(Y_i)$: Le coût du sinistre individuel.

Les coûts individuels sont i.i.d., on suppose que tous les cas de sinistres sont indépendants. Dans le cas où la fréquence et les charges sont hétérogènes. Selon des caractéristiques différant dans la base de la fréquence (variables quantitatives et qualitatives) Information, la prime pure égale :

$$E(S| \text{Information}) = E(N| \text{Information}) * E(Y_i| \text{Information}).$$

Le facteur « Information » est hétérogène, on ne peut pas comprendre directement. On pourrait utiliser des variables tarifaires dans notre base de données pour obtenir des espérances conditionnelles approché. On cherche alors A = ($A_{Grand\ sinistre}$, $A_{Moyen\ sinistre}$, $A_{Petit\ sinistre}$), c'est un ensemble de variables explicatives :

$$E(S|A) = E(N|A) * E(Y_i|A).$$

III. Modélisation de la fréquence des sinistres

Dans notre base de données relative à la fréquence des sinistres, nous disposons de trois variables quantitatives et de quatre variables qualitatives.
Nous avons calculé la probabilité d'exposition aux sinistres, et les résultats obtenus sont illustrés dans l'annexe 1, qui présente les probabilités de survenance de sinistres en fonction de différentes caractéristiques telles que l'âge du conducteur, l'âge du véhicule, la puissance du moteur (*power*), etc.

Afin d'améliorer la pertinence du modèle, nous procédons à une segmentation ainsi qu'à une sélection des variables selon leur nature et leur influence sur la fréquence des sinistres, comme indiqué ci-dessous :

- **Segmentation et sélection des variables quantitatives**

Nom des variables	Segmentation des moyennes sinistres	Réduction et réservation de l'intervalle de variable
Age du conducteur[1]	[18,20],[20,23),(23,32),[32,52),[52,84),[84,Infini)	[32,52)
Age du véhicule[2]	[0,18],[18,27),[27,Infini)	[0,18)
Density[3]	[2,363),[363,1663),[1633,Inf)	[2,363)

➢ **Segmentation des variables qualitatives :**

[1] Dans notre rapport, la caractéristique de « Age du conducteur », on a toujours choisi le « [32,52)», même niveau.

[2]Dans notre rapport, la caractéristique de « Age du véhicule », on a toujours choisi le « [0,18) », même niveau.

[3]Dans notre rapport, la caractéristique de « Density », on a toujours choisi le « [2,363)», même niveau.

Nom des variables	Segmentation des moyennes sinistres	Réduction et réservation de l'intervalle de variable
Power	"d"= "P1" "e","f","l","g","h" = "P2" "m","o","k","j","i"= "P3" "n" = "P4"	"e","f","l","g","h" = "P2"
Région	"Centre","Basse-Normandie","Bretagne","Haute-Normandie" = "R1" "Aquitaine","Pays-de-la-Loire","Poitou-Charentes" = "R2" "Limousin ","Nord-Pas-de-Calais" = "R3" "Ile-de-France" = "R4"	"Aquitaine","Pays-de-la-Loire","Poitou-Charentes" = "R2"
Brand	"Japanese (except Nissan) or Korean","Renault, Nissan or Citroen" = "V1" "Fiat","other" = "v2" "Mercedes, Chrysler or BMW","Opel, General Motors or Ford","Volkswagen, Audi, Skoda or Seat" = "v3"	"Fiat","other" = "v2"
Gas[4] (Mode de fonctionnement)	Diesel Regular	Regular

- La régression du modèle GLM

Dans notre modèle de fréquence, basé sur un **GLM suivant la loi de Poisson**, nous avons constaté que certaines variables présentaient des **p-values non significatives**. En particulier,

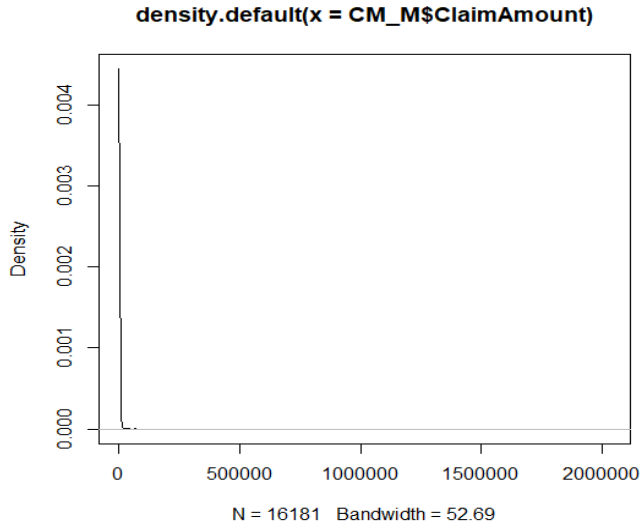
[4] Dans notre rapport, la caractéristique de Gas, on a toujours choisi le « Regular », même type.

la variable **« Région »** n'avait pas d'effet statistiquement significatif sur la fréquence des sinistres.
Par conséquent, nous l'avons supprimée du modèle, ce qui a permis d'obtenir des **résultats plus cohérents et un meilleur ajustement global**.

IV. Modélisation du coût des sinistres

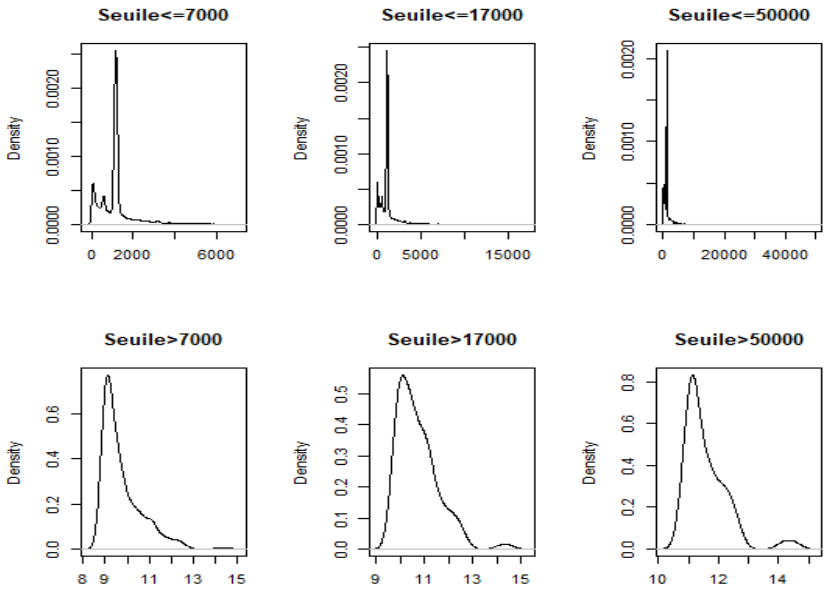
Nos deux bases de données possèdent une clé primaire commune, intitulée « PolicyID ». Grâce à cette clé, il est possible de fusionner les deux bases à l'aide de la fonction merge() dans le logiciel R, afin d'obtenir une base complète regroupant à la fois les informations sur la fréquence et le coût des sinistres.

La variable représentant le coût des sinistres est présentée ci-dessous :



En observant ce graphique, il apparaît clairement que la distribution du coût des sinistres **ne suit pas une loi Gamma**. La forme du graphique ne correspond pas à celle attendue pour une telle loi.
Afin de mieux modéliser le comportement des coûts élevés, nous avons utilisé les packages **tea** et **eva** dans **R**, qui permettent de **déterminer un seuil de coût** au-delà duquel les sinistres peuvent être considérés comme extrêmes.

Après cette étape, nous avons obtenu le graphique suivant :




```
> Res1
$second.order.par
[1]  1.018331 -0.752034

$k0
[1] 442

$threshold
[1] 6963

$tail.index
[1] 1.131785

> Res2
$sec.order.par
[1] -5.840086

$k0
[1] 61

$threshold
[1] 49088

$tail.index
[1] 1.220967

> Res3
$k0
[1] 148

$threshold
[1] 17307

$tail.index
[1] 0.9878198
```

- **Création de trois bases de données**

À partir des résultats obtenus, nous avons pu déterminer le **seuil de coût des sinistres** dans notre base de données.
Dans ce rapport, nous analysons trois **niveaux de coût de sinistre** distincts :

- **Niveau faible** : coût du sinistre inférieur à **950** ;
- **Niveau moyen** : coût du sinistre compris entre **950 et 17 000** ;
- **Niveau élevé** : coût du sinistre supérieur à **17 000**.

Ainsi, nous avons constitué **trois nouvelles bases de données**, chacune correspondant à un niveau de coût spécifique.

- **Modélisation GLM (fréquence et coût du sinistre) selon les différentes bases de données**

Comme pour l'analyse de la fréquence des sinistres, nous avons d'abord procédé à la **segmentation** et à la **sélection des variables** pertinentes.
Ensuite, nous avons effectué la **régression du modèle GLM** en supposant une **loi de Gamma** pour le coût des sinistres.
Les variables dont les **p-values n'étaient pas significatives** ont été supprimées du modèle afin d'améliorer la qualité de l'ajustement.

- **Calcul de la prédiction pour la fréquence et le coût du sinistre**

Nous utilisons les **équations de prédiction** issues des modèles pour estimer l'évolution de la **fréquence** et du **coût des sinistres**.
En multipliant la **prédiction de la fréquence** par la **prédiction du coût**, nous obtenons la **prime pure** pour chacune des bases de données partielles.

Les résultats de la segmentation varient selon la base de données considérée. Ainsi, de nouvelles **segmentations** ont été effectuées à partir des observations issues des graphiques.
Dans notre rapport, nous avons particulièrement **révisé les segmentations et la sélection** des variables qualitatives suivantes : **« power »**, **« brand »** et **« région »**.
Toutes les autres variables sont restées inchangées.

À chaque suppression d'une variable dans le modèle de régression GLM, la **valeur de l'AIC** a augmenté, indiquant une **perte d'information** dans le modèle.
Pour les **niveaux de coût élevés** et **faibles**, plusieurs variables ont dû être retirées, ce qui a entraîné une diminution de la qualité d'ajustement en raison d'une **réduction du volume de données exploitables**.

Enfin, il est important de noter que **plus le coût du sinistre est extrême** — qu'il soit très élevé ou très faible, **plus la marque du véhicule** et **la région** jouent un rôle déterminant dans la modélisation du risque.

Moyen niveau du coût de sinistre		Petit niveau du coût de sinistre		Grand niveau du coût de sinistre	
Fréquence GLM	Coût GLM	Fréquence GLM	Coût GLM	Fréquence GLM	Coût GLM
Age de conducteur	Age de conducteur	Age de conducteur	Age de conducteur	Age de conducteur	Age de conducteur
Age de véhicule	Age de véhicule	Age de véhicule	Age de véhicule	Age de véhicule	Age de véhicule
Density	Density	Density	Density	Density	Density
Power	Power	Power	Power	Power	Power
Région	Région	Région	Région	Région	Région
Brand	Brand	Brand	Brand	Brand	Brand
Gas	Gas	Gas	Gas	Gas	Gas

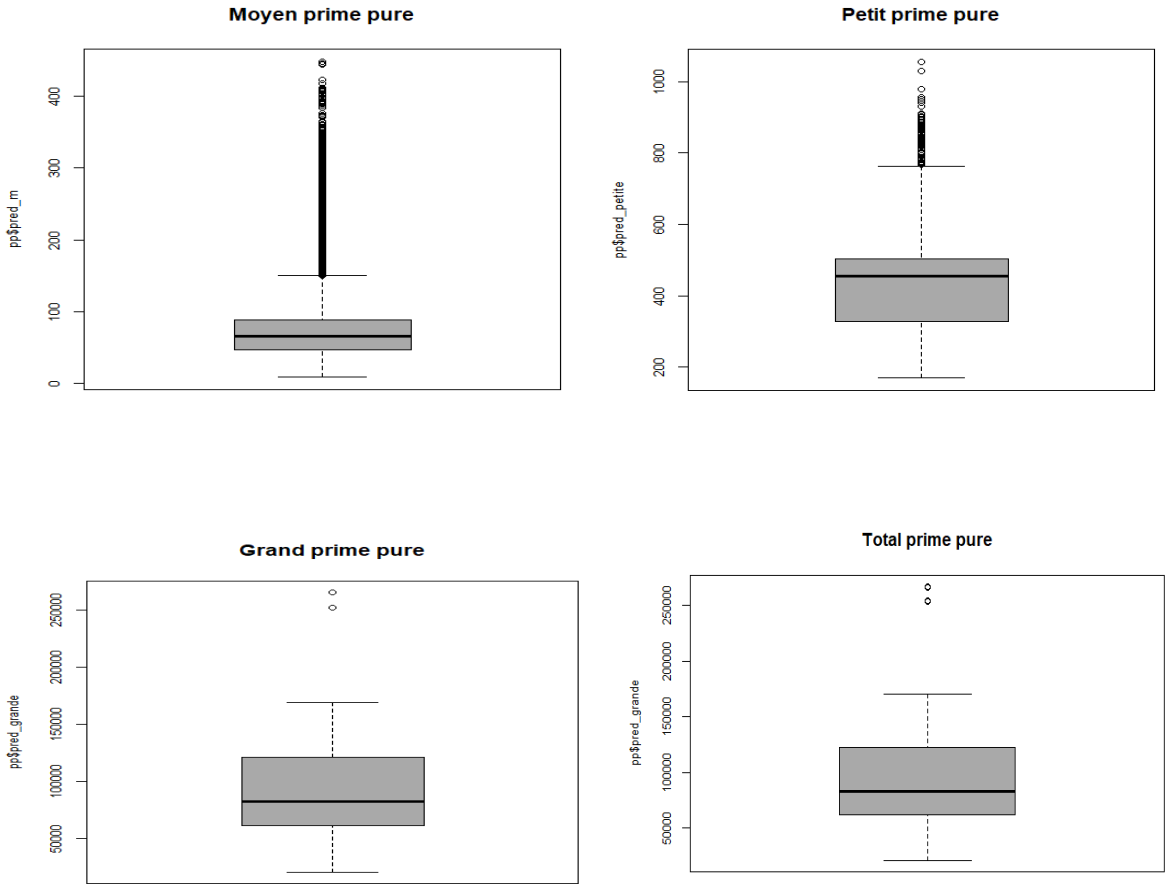
[5]

V. Détermination de la prime pure.

Prime pure	Minimum	1ère Qu	Médian	Mean	3ème Qu	Maximum
Moyen niveau du coût de sinistre générale	9,288	46,355	65,384	69,841	88,308	450,872
Petit niveau du coût de sinistre	173,2	327,4	461,6	432,8	505,9	1007,5
Grand niveau du coût de sinistre	20728	61627	82522	93290	121643	265701
Total niveau du coût de sinistre	20944	61913	83102	93792	122172	266793

[5] **Remarque importante :** les **fréquences observées** varient entre les trois bases de données, ce qui influence la qualité de l'ajustement du modèle.
Les **zones ombrées** indiquent les **observations exclues** du modèle GLM, soit en raison de valeurs extrêmes, soit de données manquantes.

Nous avons obtenu les **graphiques quantiles** présentés ci-dessous :



VI. Analyse des résultats et conclusion

En comparant les trois niveaux de coût des sinistres, on constate que la **prime pure dans la base « Petit niveau du coût de sinistre »** est inférieure à celle observée dans la base « Grand niveau du coût de sinistre ».

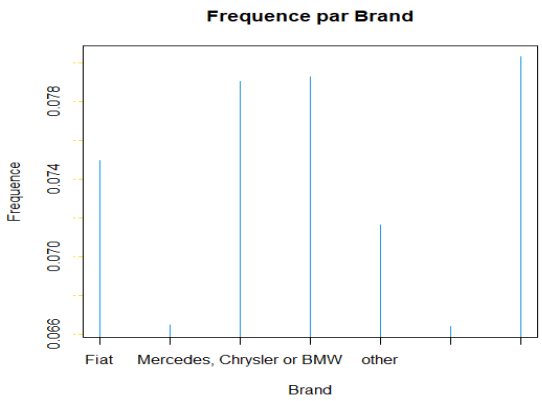
Cependant, le calcul de la **prime pure totale** présenté dans notre rapport n'est pas totalement représentatif. En effet, la fréquence des sinistres dans la base « Moyen niveau du coût de sinistre » est très faible, car cette base correspond à des données générales. Cette faible fréquence entraîne une **prime pure plus faible** que pour les autres niveaux.

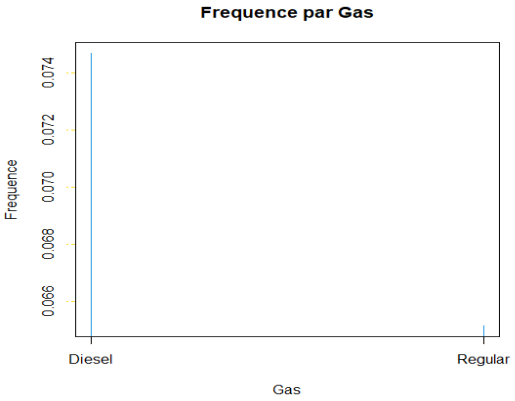
En observant le graphique relatif au **« Moyen prime pure »**, on remarque la présence de **valeurs extrêmes** : de nombreux cas présentent des primes pures supérieures à la moyenne de ce niveau. En revanche, le graphique du **« Grand prime pure »** montre que les estimations sont plus concentrées et donc **plus précises** pour cette catégorie.

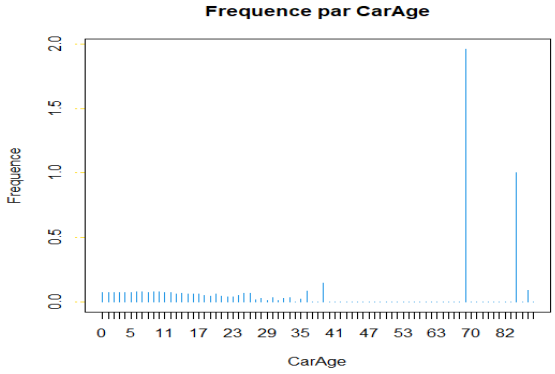
- ## Annexes

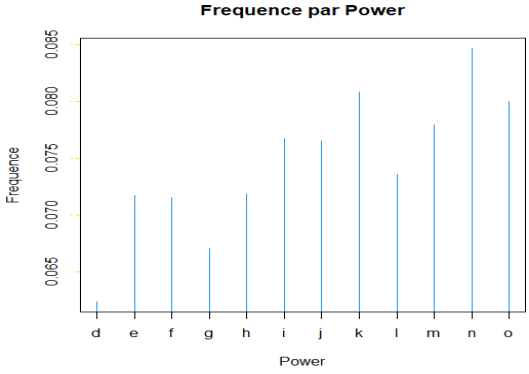
Annexe 1 :

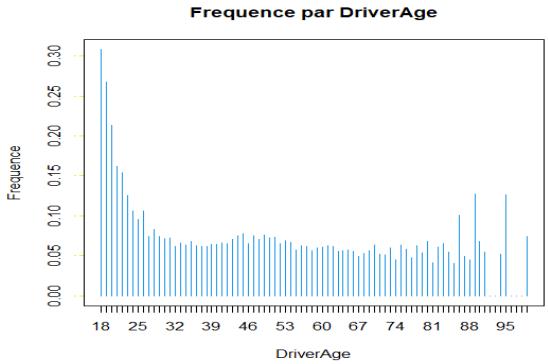
Segmentation des informations différents :

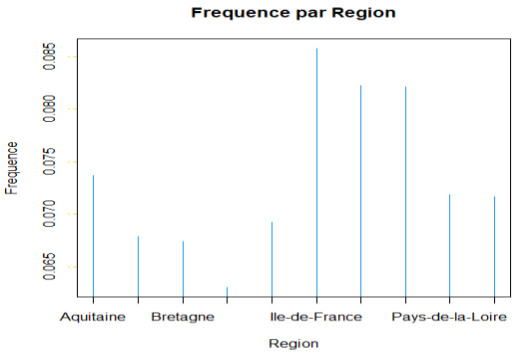


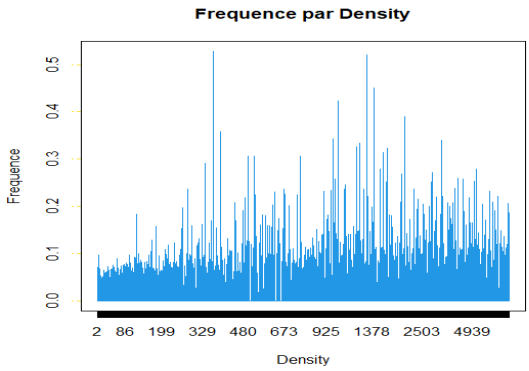












Annexe 2 :

Résultat de la régression de la fréquence GLM dans la base de données complète :

```
Call:
glm(formula = ClaimNb ~ Power_m + Brand_m + Region_m + DriveAge_m +
    CarAge_m + Density_m + Gas_m + offset(Exposure), family = poisson(link = "log"),
    data = FREQ.B)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-0.7061  -0.3109  -0.2652  -0.2230   5.8170

Coefficients:
                          Estimate Std. Error  z value Pr(>|z|)
(Intercept)             -4.011506   0.037494 -106.990  < 2e-16 ***
Power_mP1               -0.106643   0.023385   -4.560 5.11e-06 ***
Power_mP3                0.075195   0.024897    3.020 0.002526 **
Power_mP4                0.132687   0.134459    0.987 0.323729
Brand_mV1               -0.110907   0.031395   -3.533 0.000411 ***
Brand_mv3                0.031950   0.033829    0.944 0.344932
Region_mLimousin         0.223198   0.073469    3.038 0.002382 **
Region_mR1              -0.002735   0.020325   -0.135 0.892964
Region_mR3              -0.088562   0.037159   -2.383 0.017157 *
Region_mR4              -0.035289   0.028478   -1.239 0.215279
DriveAge_m[18,20]        1.125895   0.049470   22.759  < 2e-16 ***
DriveAge_m(20,23]        0.648749   0.040192   16.141  < 2e-16 ***
DriveAge_m(23,32]        0.055648   0.022528    2.470 0.013504 *
DriveAge_m(52,84]       -0.050729   0.018770   -2.703 0.006877 **
DriveAge_m(84,Inf]       0.234973   0.117137    2.006 0.044859 *
CarAge_m(18,27]         -0.305618   0.057483   -5.317 1.06e-07 ***
CarAge_m(27,Inf]        -0.921577   0.208826   -4.413 1.02e-05 ***
Density_m(363,1.66e+03]  0.252668   0.019408   13.019  < 2e-16 ***
Density_m(1.66e+03,Inf]  0.323996   0.023421   13.834  < 2e-16 ***
Gas_mDiesel              0.139396   0.016623    8.386  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for poisson family taken to be 1)
```

```
(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 104504  on 413168  degrees of freedom
Residual deviance: 103315  on 413149  degrees of freedom
AIC: 134613
```

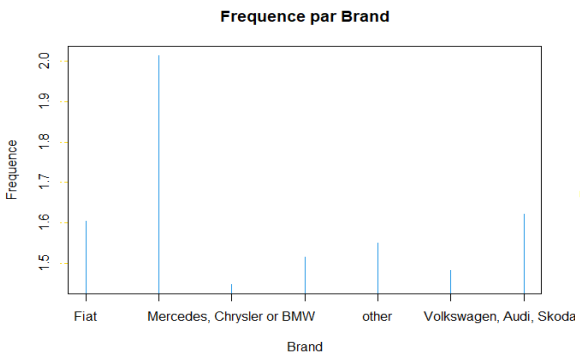
- Réduction la variable « Région »

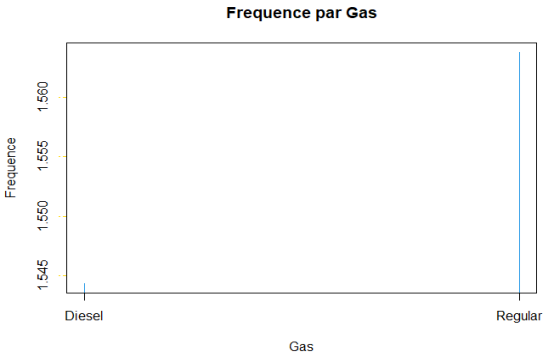
```
Call:
glm(formula = ClaimNb ~ Power_m + Brand_m + DriveAge_m + CarAge_m +
    Density_m + Gas_m + offset(Exposure), family = poisson(link = "log"),
    data = FREQ.B)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-0.7024  -0.3111  -0.2655  -0.2236   5.7751

Coefficients:
                       Estimate Std. Error  z value Pr(>|z|)
(Intercept)            -4.01022    0.03416 -117.406  < 2e-16 ***
Power_mP1              -0.10773    0.02337   -4.610 4.02e-06 ***
Power_mP3               0.07456    0.02486    2.999 0.002708 **
Power_mP4               0.13036    0.13445    0.970 0.332255
Brand_mV1              -0.11390    0.03136   -3.632 0.000281 ***
Brand_mv3               0.03057    0.03382    0.904 0.366036
DriveAge_m[18,20]       1.12638    0.04944   22.785  < 2e-16 ***
DriveAge_m(20,23]       0.64905    0.04016   16.161  < 2e-16 ***
DriveAge_m(23,32]       0.05327    0.02250    2.367 0.017923 *
DriveAge_m(52,84]      -0.04896    0.01874   -2.612 0.008997 **
DriveAge_m(84,Inf]      0.23720    0.11711    2.026 0.042813 *
CarAge_m(18,27]        -0.30182    0.05746   -5.253 1.50e-07 ***
CarAge_m(27,Inf]       -0.92184    0.20882   -4.415 1.01e-05 ***
Density_m(363,1.66e+03]  0.24069  . 0.01902   12.652  < 2e-16 ***
Density_m(1.66e+03,Inf]  0.30162    0.01980   15.234  < 2e-16 ***
Gas_mDiesel             0.13849    0.01659    8.346  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 104504  on 413168  degrees of freedom
Residual deviance: 103332  on 413153  degrees of freedom
AIC: 134622
```

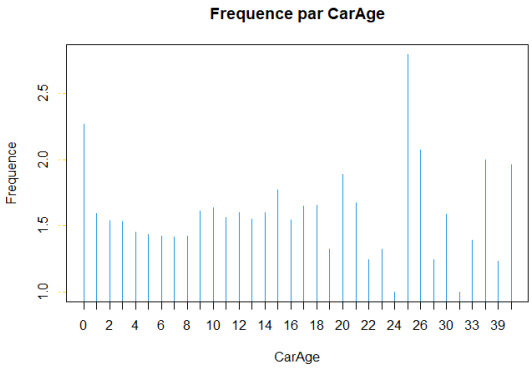
```
(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 104504  on 413168  degrees of freedom
Residual deviance: 103332  on 413153  degrees of freedom
AIC: 134622

Number of Fisher Scoring iterations: 6
```

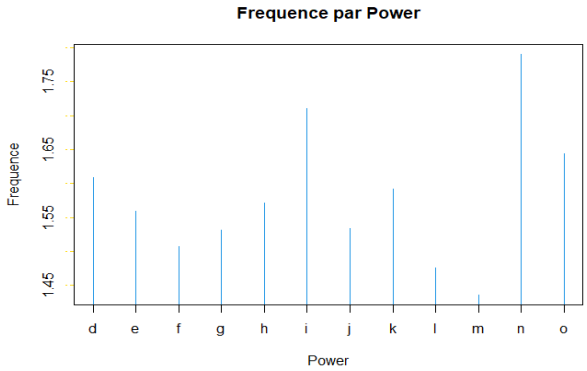
Annexe 3 :

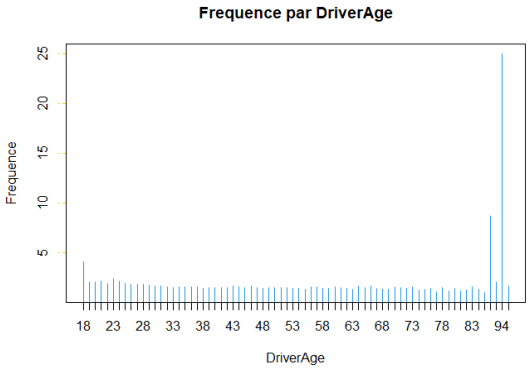
- Petit niveau du coût de sinistre :

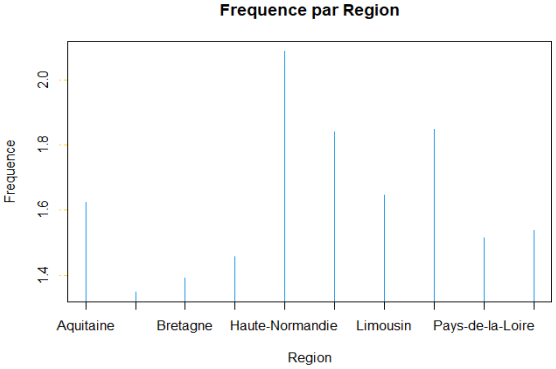

Frequence par Brand
Frequence
1.5
1.6
1.7
1.8
1.9
2.0
Fiat
Mercedes, Chrysler or BMW
other
Volkswagen, Audi, Skoda
Brand

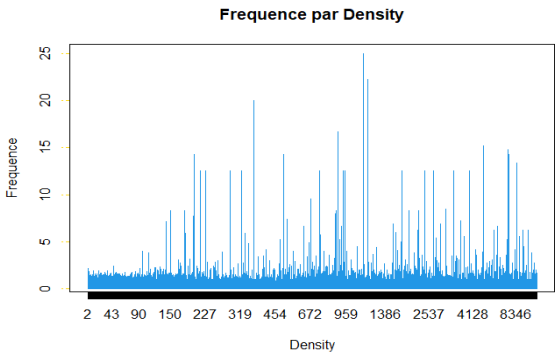

Frequence par Gas
Frequence
1.545
1.550
1.555
1.560
Diesel
Regular
Gas


Frequence par CarAge
Frequence
1.0
1.5
2.0
2.5
0 2 4 6 8 10 12 14 16 18 20 22 24 26 30 33 39
CarAge

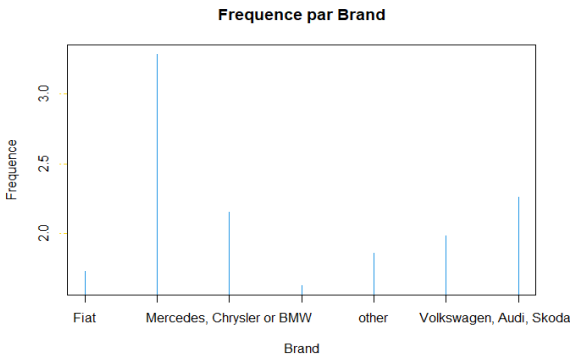

Frequence par Power
Frequence
1.45
1.55
1.65
1.75
d e f g h i j k l m n o
Power

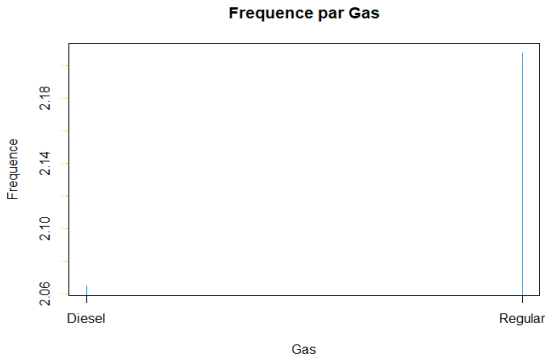

Frequence par DriverAge
Frequence
5
10
15
20
25
18 23 28 33 38 43 48 53 58 63 68 73 78 83 94
DriverAge

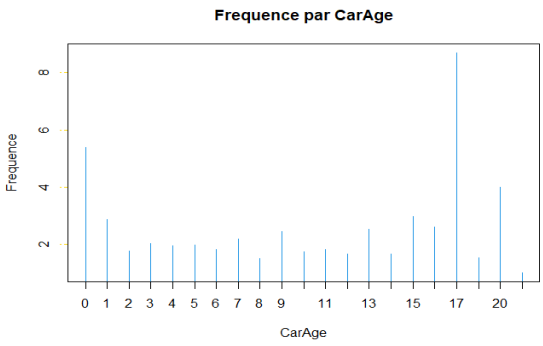

Frequence par Region
Frequence
1.4
1.6
1.8
2.0
Aquitaine
Bretagne
Haute-Normandie
Limousin
Pays-de-la-Loire
Region

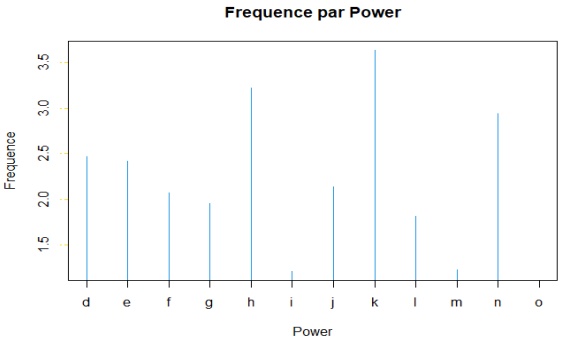


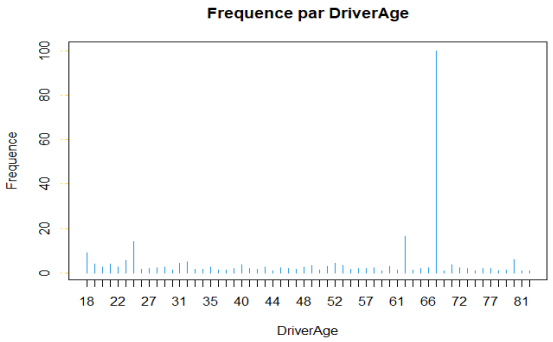
- Grand niveau du coût de sinistre :

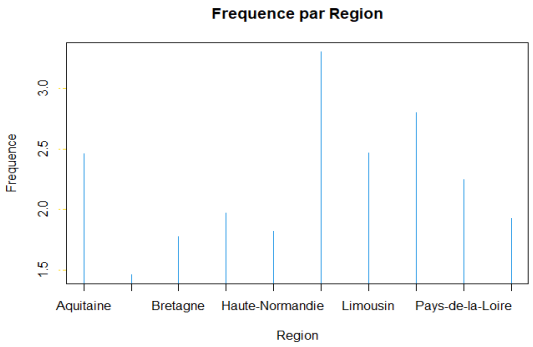


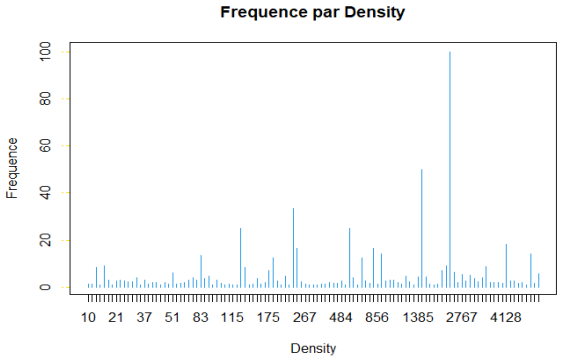







Frequence par DriverAge
Frequence
0
20
40
60
80
100
18 22 27 31 35 40 44 48 52 57 61 66 72 77 81
DriverAge


Frequence par Region
Frequence
1.5
2.0
2.5
3.0
Aquitaine
Bretagne
Haute-Normandie
Limousin
Pays-de-la-Loire
Region


Frequence par Density
Frequence
0
20
40
60
80
100
10 21 37 51 83 115 175 267 484 856 1385 2767 4128
Density